OPEN ACCESS

University of BRISTOL

Roberts, L., McCahon, D., Johnson, O., Haque, M. S., Parle, J., & Hobbs, F. D. R. (2018). Stability of thyroid function in older adults: The Birmingham Elderly Thyroid Study. *British Journal of General Practice*, *68*(675), e718-e726. https://doi.org/10.3399/bjgp18X698861

Peer reviewed version

Link to published version (if available):
10.3399/bjgp18X698861

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at https://onlinelibrary.wiley.com/doi/full/10.1111/bjdp.12253 . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

Title: Stability of thyroid function in older adults: the Birmingham Elderly Thyroid Study

Short running title: Stability of thyroid function in older age

Authors:

1) Professor Lesley Roberts, PhD, The Medical School, University of Warwick, UK
2) Dr Deborah McCahon, PhD, Centre for Academic Primary Care, University of Bristol, UK
3) Dr Oliver Johnson, MBChB, Countess of Chester Hospital, Chester, UK
4) Dr M.Sayeed Haque, PhD, Institute of Applied Health Research, University of Birmingham, UK
5) Professor James Parle, MD, Institute of Clinical Sciences, University of Birmingham, UK
6) Professor FD Richard Hobbs, FMedSci, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

*Corresponding author: Professor FD Richard Hobbs, Nuffield Department of Primary Care Health Sciences, Radcliffe Primary Care Building, Radcliffe Observatory Quarter, Woodstock Road, University of Oxford, Oxford, UK. Email: Richard.hobbs@phc.ox.ac.uk Tel: 1865 617851

Word count 2552

Abstract

Background: Thyroid function tests (TFT) are among the most requested tests internationally. Testing practice is however inconsistent and potentially sub-optimal and overly costly. The natural history of thyroid function remains poorly understood.

Aim: To establish the stability of thyroid function over time, and identify predictors of development of overt thyroid dysfunction
Design: Longitudinal follow-up

Setting: 19 general practices**,** UK

Method: 2936 Birmingham Elderly Thyroid Study (BETS I) with a baseline TFT result indicating euthyroid or subclinical state were re-tested after approximately 5 years. Change in TSH, FT_4 and thyroid status between baseline, and follow-up was determined. Predictors of progression to overt dysfunction were modelled.

Results: Participants contributed 12,919-person years, 17 cases of overt thyroid dysfunction were identified, 13 having been classified at baseline as euthyroid and four as having subclinical thyroid dysfunction. Individuals with subclinical results at baseline were 10 and 16-fold more likely to develop overt hypo and hyperthyroidism respectively compared with euthyroid individuals. TSH and FT_4 demonstrated significant stability over time with 61% of participants having a repeat TSH concentration within 0.5 mIU/L of their original result. Predictors of overt hypothyroidism included new treatment with amiodarone (OR 92.1), a new diagnosis of atrial fibrillation (OR 7.4), or renal disease (OR 4.8).

Conclusion: High stability of thyroid function demonstrated over the 5-year interval period should discourage repeat testing, especially where a euthyroid result is in the recent clinical record. Reduced repeat TFTs in older individuals is possible without conferring risk and could result in significant cost savings.

Key terms: Thyroid function test, Subclinical thyroid dysfunction, symptoms, aging, primary health care, general practice

How this fits in

- Thyroid tests are commonly requested in the routine care of older adults.
- Subclinical thyroid dysfunction is a relatively common biochemical finding among elderly patients.
- Current practice for the management of a single test indicating mildly abnormal thyroid dysfunction in the elderly population is variable and potentially sub-optimal and overly costly.
- This large, population-based survey demonstrates significant stability in thyroid function over a period of up to 5 years in the elderly population, with 96% of individuals who were euthyroid at baseline remaining so and less than 0.5% developing an overt hyper or hypothyroid status.
- Based upon this evidence, routine repeat thyroid function testing among elderly individuals who have a recent (within 5 years) euthyroid result in their clinical record is not advised, unless clinically indicated.

Introduction

Disorders of the thyroid include both overt and subclinical hypothyroidism and hyperthyroidism. Subclinical thyroid dysfunction is common among the elderly, characterised by serum thyroid stimulating hormone (TSH) outside the reference range in association with serum thyroid hormone (free thyroxine (FT_4) and triiodothyronine (FT_3)) concentrations within the reference range.[1-3] Both subclinical thyroid states are of limited clinical relevance. In overt thyroid disease states, clinicians are generally more concerned about hypothyroidism since onset is often non-specific and insidious, so the diagnosis is often missed. In contrast, hyperthyroidism will normally present with less common symptoms and be diagnosed promptly.

Within the UK, systematic screening is not recommended, partly because 'the natural history of thyroid function is unclear'.[4] Since there is also uncertainty over management of subclinical disease, current practice is both variable and potentially sub-optimal or excessively costly.[5–7] Recommendations for how often thyroid function tests (TFTs) should be repeated after a previously normal or subclinical test result are lacking. The annual estimated cost of TFTs in the UK is £30 million with the majority originating in primary care.[5, 8–9] Recent work undertaken by the authors suggests that annually in UK general practice, TFTs are requested for around 30% of older patients without overt thyroid dysfunction (unpublished data). Available evidence suggests that Primary Care Physicians (PCPs) repeatedly request TFTs in this patient group, in response to vague symptoms, previously mildly abnormal tests or as part of other routine care monitoring. [8] Few studies have explored the natural history of thyroid dysfunction or the value of a single TFT within a primary care population. One small single-site primary care based study followed 73 patients with subclinical hypothyroidism for 12 months, reporting that 17.8% developed overt disease and 5.5% reverted to a euthyroid state.[1] Follow-up of the Whickham cohort identified increased odds of development of overt disease if an elevated TSH had previously been reported, but the 20-year interval used in this study and all-age cohort make application of findings to an elderly population difficult.[10]

Secondary care studies are conflicting with one study reporting an incidence of 9.9 overt cases per 100 patient years and a study with longer follow up suggesting an annual incidence of 5.6, although the female-only population derived from secondary care limits generalisability.[11,12] A further female only cohort study of 252 individuals referred for elevated TSH reported a similar progression with 19% requiring treatment for overt dysfunction or persistent elevated TSH (>10 mIU/l) over a 5 year interval[13]. This Brazilian study was conducted in a region of iodine intake inadequacy and relevance to the UK may be lacking. Whilst both populations and biochemical definitions of thyroid dysfunction differ, the much lower annual incidence reported in secondary care populations compared to the primary care study indicates more evidence is needed to identify predictors of overt disease outside of the secondary care setting.

Evidence for the progression of subclinical hyperthyroidism is similarly lacking. Sawin et al[14] reported that none of the 33 subclinically hyperthyroid patients they followed up for 4 years developed overt disease and similar findings are reported by Woeber[15] with only 1 of 16 patients followed developing overt disease. Whilst these findings demonstrate consistency, numbers are small and a recent expert panel review concluded there was insufficient evidence to comment on the natural history of subclinical hyperthyroidism.[5]

This study aimed to address some of the important gaps in the evidence base, namely to (i) determine incidence of subclinical thyroid dysfunction in an elderly primary care population previously shown to be euthyroid (ii) establish the proportion of patients with subclinical thyroid dysfunction who revert to a euthyroid state, experience persisting subclinical dysfunction or develop overt disease (iii) evaluate within-person variation in TSH and FT_4 concentrations over a 5 year interval, and (iv) identify predictors of progression to overt hyper/hypothyroidism.

Methods

Background to the Birmingham Elderly Thyroid Study (BETS I)

This study comprises follow-up of the BETS I cohort, a screening study of 5881 patients aged over 65 from 20 practices representative of the UK, conducted between 2002 and 2004. [16-19] Thyroid function tests comprised measurement of TSH and FT_4. Measurement of FT_3 was undertaken as dictated by routine laboratory protocol.

Practice recruitment

19 of the original 20 practices agreed to participate.

Inclusion criteria

- Patients identified in BETS I as having thyroid function results within normal or subclinical ranges. A euthryroid status was defined as both TSH and FT_4 being within ranges indicated in Table 1. Subclinical hypothyroidism and hyperthyroidism were defined as FT_4 in range and TSH being above or below reference range respectively.

Exclusion criteria

BETS I participants

- for whom classification of thyroid status was not possible, or who had overt thyroid dysfunction at baseline
- who were recruited to the active treatment arm of the RCTembedded within BETS 1[17]
- for whom the responsible clinician deemed contact inappropriate
- unable or unwilling to give informed consent

Study procedure

All eligible patients were sent an invitation letter, patient information sheet, and response return slip, after receipt of which screening appointments were organised at their usual surgery. TFTs for BETS 2 occurred over a 9-month period, approximately 5 years after the initial screening.

Case note evaluation

Data on diagnoses and treatment including known confounders such as amiodarone (which comprises 37% iodine) were collected from primary care records. All significant medical diagnoses were categorised in accordance with recognised major disease groups as reported in BETS I.[16]

For participants having thyroid surgery, radio-iodine therapy or starting anti-thyroid drugs, and thyroxine replacement therapy in the interval period, the results of the TFTs immediately prior to commencement of treatment were extracted. If one or both measurements had not been conducted immediately prior to initiation of treatment, medical records were reviewed to ascertain reason for commencement of therapy and enable classification of thyroid status.

Measurement and categorisation of thyroid function

TFTs were measured by electrochemiluminescent immunoassays (Roche E170, Roche Diagnostics, UK). The TSH assay was calibrated against the second International Reference Preparation 80/558 (lower limit of reporting; 0·02 mIU/L, manufacturer's quoted mean functional sensitivity; 0·014 mIU/L). Laboratory reference ranges are reported in Table 1. Changes in the assays used to measure TSH, FT_4 and FT_3 occurred between the two studies.[16] In parallel with this, a change to the standard reference ranges for TSH, FT_4, and FT_3 occurred. To enable comparison across the two-time points, a correction factor was applied to the baseline TFT results and subsequent reclassification of thyroid status was undertaken before data were compared across the time interval.

Thyroid status was classified as euthyroid, subclinically hypothyroid, overtly hypothyroid, subclinically hyperthyroid or overtly hyperthyroid based upon the reference ranges reported in Table 1.

Follow-up contribution in terms of patient years at risk was calculated for all subjects as the time interval between initial and follow-up screen. Those receiving thyroid function treatment were classified based upon the results of the TFT immediately prior to commencement of treatment and their contribution censored at this point.

Primary analysis

Incidence was calculated as number of cases divided by number of person-years at risk. Risks of developing disease were calculated, and risks compared for groups who were categorised as euthyroid and subclinical at BETS I. Sensitivity analysis was performed reclassifying all patients who commenced therapy during the interval period as having overt thyroid dysfunction (e.g. overt hyper or hypothyroidism based upon treatment given) and re-running analyses.

Binary logistic regressions were performed to identify predictors of development of overt hyper/hypothyroidism. The forward stepwise LR method was used to identify variables with a significance level of 5% for inclusion and 10% for removal. Two additional variables were created, one to indicate subclinical thyroid status at BETS I and the other to indicate whether BETS I thyroid function status had been reclassified due to the application of the laboratory defined correction factors. In total, 31 variables were available for construction of r models including

medical conditions (classed as absent/pre-existing/new (occurring in the interval between studies)), amiodarone use, alcohol intake, smoking status, family history of thyroid dysfunction, age, and deprivation measure (IMD 2004 score).[20]
Analyses were undertaking using SPSS 15.0.

Exploratory analysis

To explore stability of TSH and FT_4 over time, change in both measures between BETS I and BETS 2 was calculated for each participant and BETS I values plotted against BETS 2 values to explore within-person change.

Results

Eligibility for follow-up

Overall 103 BETS I participants were ineligible for follow-up *(Figure 1).* A further 1335 were deceased or excluded prior to invitation. Just under 50% of the BETS I cohort (n=2936) fulfilled the inclusion criteria and attended for follow-up. Those available to follow-up demonstrated a statistically significant difference in deprivation scores and age compared to those not available to follow-up. The screened cohort had baseline IMD scores indicative of marginal greater affluence (mean IMD 21.65 versus 25.65) and were on average 1.82 years younger. This difference is likely to be attributable to greater mortality in older and less affluent participants, but given the large sample and 5 year follow-up period of an older adult cohort these difference are unlikely to impact findings.

Population characteristics

Age ranged from 68·7–96·4 years with a mean of 76·9 years (standard deviation 5·03) and 49% were female. Socioeconomic status ranged from 3·16 (most affluent) to 74·4 (least affluent), mean IMD score; 21·79 (SD 15·14).

Overall 92·3% (2709/2936) were classified as euthyroid, 1% (n=29) subclinically hyperthyroid, 6·2% (n=181) subclinically hypothyroid, 0·3% (n=8) were classified as overtly hyperthyroid and 0·3% (n=9) overtly hypothyroid. 1·8% (53/2396) had a thyroid diagnosis or treatment in the interim period and were classified based upon the TFT immediately prior to treatment commencement.

Status change over time

Overall, 95.5% (95% CI 94.7, 96.3%, 2644/2768) of the individuals classified as euthyroid at baseline retained euthyroid status at follow-up. Six (0·2%, 95%CI 0.1, 0.5%) classified as euthyroid at baseline had developed overt hypothyroidism, and seven (0·3%, 95% CI 0.1, 0.5%) had developed overt hyperthyroidism. Only 3·5% (95%CI 2.9, 4.3%, 98/2768) had follow-up results indicative of a change from euthyroid to subclinical hypothyroidism, and 0·5% (95% CI 0.3, 0.8%), 13/2768to subclinical hyperthyroidism (Table 2).

Of the 25 individuals classified as subclinically hyperthyroid at baseline, 16 (64%, 95%CI 42.5, 82.0%) retained this classification, eight (32%, 95%CI 15.0, 53.5%) reverted to euthyroid status, and one (4%, 95% CI 0.1, 20.4%) individual developed overt hyperthyroidism. A similar proportion, 58% (95%CI 49.5, 66.2%, n=83/143), classified as being subclinically hypothyroid at baseline remained so, 40% (95% CI 31.8, 48.4%, n=57) reverted to euthyroid status whilst 2% (95%CI 0.4, 6.0%, n=3) developed overt hypothyroidism.

Sensitivity analysis

Since therapy may have disrupted natural history and dysfunction progression, all treated cases were re-categorised as having overt thyroid dysfunction. The total number of cases of overt hypothyroidism assumed therefore increases to 51 (26 being euthyroid and 25 subclinical at baseline), suggesting that a maximum of 1·7% may have developed overt hypothyroidism compared to the 0·3% estimate based upon TFT results alone. The total number of cases of overt hyperthyroidism remains the same because in all cases therapy was commenced based upon TFT results in the overt range.

Cases of overt dysfunction

Overall, 2936 participants contributed 12,919 person-years for analysis. The risk of developing overt hypothyroidism in the subclinical hypothyroid group was 51·5 per 10,000 person-years at risk (95% CI 38.8, 67.1) compared to 4·9 (95% CI 1.6, 10.2) per 10,000 person-years at risk in the euthyroid group, making the subclinical group 10 times more likely to develop overt hypothyroid dysfunction.

Sensitivity analyses increases risk for an individual with a subclinical hypothyroid status to 20 times that for a euthyroid individual 429·5 (95%CI 390.3, 471.9) versus 21·3 (95% CI 13.8, 32.1) per 10,000 person-years at risk respectively). It is noted that true risk is likely to be magnified in sensitivity analyses due to inclusion of 20 individuals who were euthyroid at baseline but treated with thyroxine during the interval period based on a subclinical result.

The risk of an individual with a subclinical hyperthyroid status developing overt hyperthyroidism was approximately 16 times greater than that of an individual with a euthyroid status 95·4 (95% CI 7.8, 116.1) versus 5·7 (95% CI 2.2, 11.7) per 10,000 person-years at risk respectively).

Predictors of development of overt thyroid dysfunction

Given the low number of events (overt thyroid disease) it was not possible to produce a robust model to predict development of overt dysfunction. However given that to achieve sufficient events for modelling, over 14 thousand individuals would require follow up (from a 28 thousand baseline population, based on our follow-up), we have provided a cautious model based on available data to identify factors which appear to increase risk of development of hypothyroidism. (Table 3). Individuals with higher TSH or lower FT_4 values at baseline are at greatest risk of development of overt hypothyroid status. Later diagnoses of atrial fibrillation or renal disease, or commencement of amiodarone also increase likelihood of progression.

The model constructed to predict development of hyperthyroidism failed to demonstrate any predictors, which is not surprising given the very low event rate..

Change in TSH and T_4

Change in TSH between BETS I and BETS 2 was calculated and explored using an arbitrary definition of a shift of ≥0·5 units being classified as 'change'. 61% of participants showed no change between BETS I and BETS 2 using this definition. Removal of outliers further demonstrates equal numbers of individuals experiencing an increase and decrease in TSH between time points, and, for the majority, TSH remains relatively stable. *(Figure 2).*

Change in FT_4 was calculated based upon a definition of change as ≥1·0 unit, 39·6% (n=1162) of participants showed no change using this definition with very few individuals demonstrating large shifts in FT_4 (*Figure 3*).

Discussion

Summary

This large population-based study provides the most comprehensive data yet on the long-term dynamics of thyroid function among elderly patients in primary care. The prevalence of subclinical hypothyroidism in this study population was 6·2%, similar to that reported in smaller studies.[21] Subclinical hyperthyroidism was less common, with just 1% of individuals being affected. BETS 2 confirms the incident findings of BETS 1, that subclinical thyroid dysfunction is a relatively common biochemical finding among elderly patients and therefore understanding the natural course of the disease is important for both clinicians and patients.[16]

This study also demonstrates significant stability of thyroid function in this aging population over a long (mean 5 year) time interval with 96% of individuals who were euthyroid at baseline remaining so and less than 0·5% developing an overt hyper or hypothyroid status.

Strengths and limitations

This is the largest study to date following natural history of thyroid functioning in an unselected primary care cohort of older adults, with over 2900 individuals followed over a 5 year interval. One limitation of this study was that thyroid status on each occasion was based upon a single sample. There are several factors that can influence the reproducibility of TFTs, including seasonality [22] and intercurrent illness [23], however, any subtle or transient change that may have occurred are unlikely to have influenced our findings in any significant fashion due to the large sample size and the likelihood of transient fluctuations impacting at both time points. The predictive model needs to be interpreted with caution given low event rates. An automatic selection process to create a set of variables with the strongest association with the outcome was used, given this low event rate, however such a process is data driven and may create a biased selection. However confidence is derived from the fact that all variables associated with development of hypothyroidism have strong biological plausibility.

Comparison with existing literature

Of those individuals classified as subclinically hypothyroid at baseline, only 2% progressed to overt hypothyroidism which is lower than has been previously reported. [24] Our study demonstrated 40% of such patients revert to euthryoid status which is in-line with the 15-65% estimate by Somwaru *et al* [25] and the majority of those with subclinical hypothyroidism at baseline retain a subclinical state despite the long screening interval. Similarly, just 4% of those who were subclinically hyperthyroid at baseline developed overt hyperthyroidism.

To further support stability over time we demonstrate here that most individuals (61%) had almost no change in TSH (≤0·5 mIU/L), with equal numbers of individuals experiencing either increasing or decreasing TSH. This contrasts with findings from previous work that suggests TSH

increases with age.[26] In a similar manner, we also demonstrate that FT_4 is remarkably stable over time in elderly subjects.

In this study development of overt hypothyroid status was significantly more likely among individuals with a new diagnosis of either AF or renal disease, or recent commencement of amiodarone. This supports a previous finding that prevalence of subclinical hypothyroidism is higher among patients with chronic renal disease. [27] These factors along with previously high-normal TSH or low-normal FT_4 increase likelihood of hypothyroid dysfunction and may represent triggers for repeat testing although further work is needed to describe any benefit accrued from such a targeted approach. The large Odds Ratio with wide conference intervals observed for commencement of amiodarone during the follow-up period (92·1; 95% CI 5·64-1501·39) should however be interpreted with caution as it is likely an artefact of the very low prevalence of amiodarone use in the study cohort.

Predictors of development of hyperthyroidism could not be determined from this study but considering the low incidence of hyperthyroidism there is nothing to support repeat testing of individuals with a normal test result within the previous 5 years to identify hyperthyroidism in the absence of substantive clinical signs and symptoms.

Implications for practice

This study confirms that older patients with subclinical thyroid dysfunction are only at small relative risk of progression to overt dysfunction, and absolute risk is very low. TFTs are remarkably stable over extended periods and repeat routine testing or screening should be avoided in this population, particularly where a previous euthyroid result is reported in the clinical record. Considering this evidence, clinicians should minimise repeat TFT, unless clinically indicated, among elderly individuals who have a recent (within 5 years) euthyroid result in their clinical record. This is an important finding given that almost a third of older patients without overt thyroid dysfunction have at least one TFT performed annually (unpublished data).

Table 1: Reference ranges for thyroid function assays with their associated intra assay co-efficient of variation.

Test	Reference Range	Intra –assay co-efficient of variation (associated range)
Thyroid Stimulating Hormone (TSH)	0·3-4·5 mIU/L	1·5% (0·5-33·0 mIU/L)
Free Thyroxine (FT_4)	10-22 pmol/L	2·0-2·5% (9·0-66.0 pmol/L)
Free Triiodothyronine (FT_3)	3·1-6·8 pmol/L	2·0-3·5 (4·0-21·0 pmol/L)

Table 2: Thyroid status at baseline and follow-up

	Follow-up (BETS 2) status (1 unclassified)					
Baseline (BETS 1) status	n=2936	Overt hypothyroid n= 9 (0·3%)	Subclinical hypothyroid n=181 (6·2%)	Euthyroid n= 2709 (92·3%)	Subclinical hyperthyroid n=29 (0·99%)	Overt hyperthyroid n= 8 (0·3%)
	Subclinical hypothyroid (n=143)	3 *(2.1%)*	83 *(58.0%)*	57 *(39.86%)*	0	0
	Euthyroid (n=2768)	6 *(0.2%)*	98 *(3.5%)*	2644 *(95.5%)*	13 *(0.5%)*	7 *(0.3%)*
	Subclinical hyperthyroid (n=25)	0	0	8 *(32.0%)*	16 *(64.0%)*	1 *(4.0%)*

Shaded cells indicate no status change over the screening interval

Table 3: Logistic regression – factors associated with development of hypothyroidism and associated likelihood

Variable	Co efficient	P-value	OR	Lower 95 % CI for OR	Upper 95 % CI for OR
Baseline TSH	0·38	0·001	1·46	1·17	1·81
Baseline FT_4	-0·86	<0·001	0·42	0·27	0·66
New* amiodarone prescription	4·61	0·001	92·1	5·64	1501·39
New* AF diagnosis	2·00	0·012	7·41	1·56	35·14
New* renal disease diagnosis	1·57	0·044	4·81	1·04	22·22

31 variables were available for construction of logistic regression models, comprising 22 disease categories, amiodarone use, alcohol use, smoking status, family history of thyroid dysfunction, age, gender, IMD score and baseline TSH and FT4. Seven disease groups and family history were removed to maximise the dataset

available for analysis were data were missing for 1% or more of the population. The forward stepwise method was used with a significance level of 5% for variable inclusion and 10% for variable removal. Eight variables were entered to the final model although existing (at baseline) diagnoses of AF, renal disease or amiodarone use did not make a significant contribution to the final model.

** 'New' relates to prescription or diagnosis occurring for the first time after the initial baseline screening episode*

Figure 1. Consort diagram

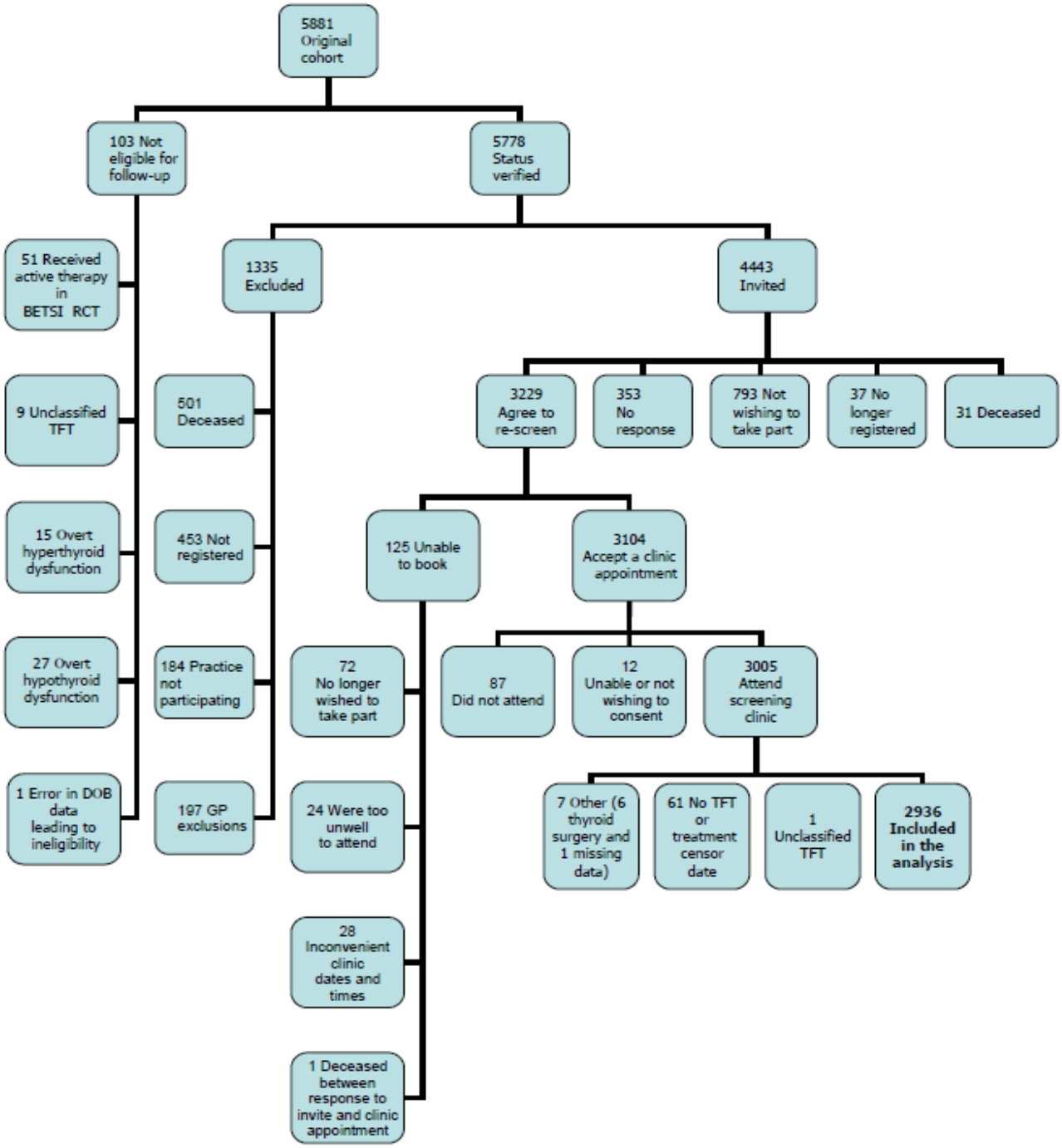
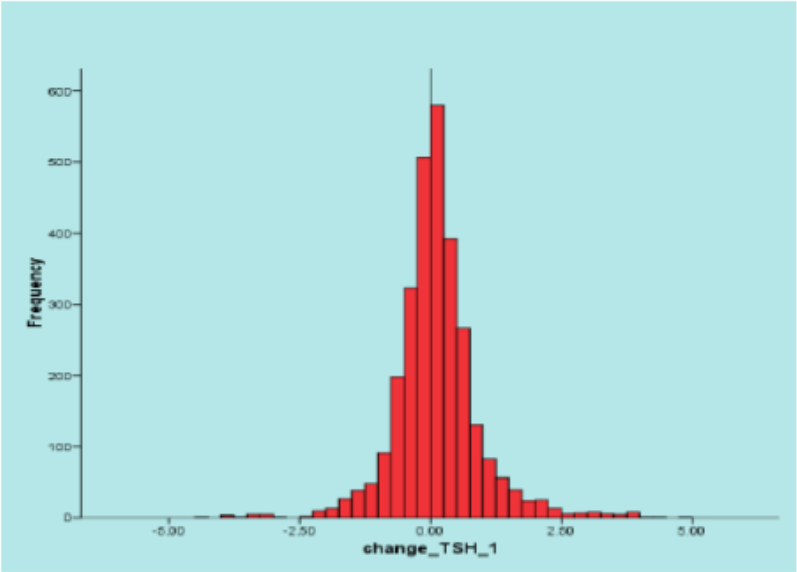
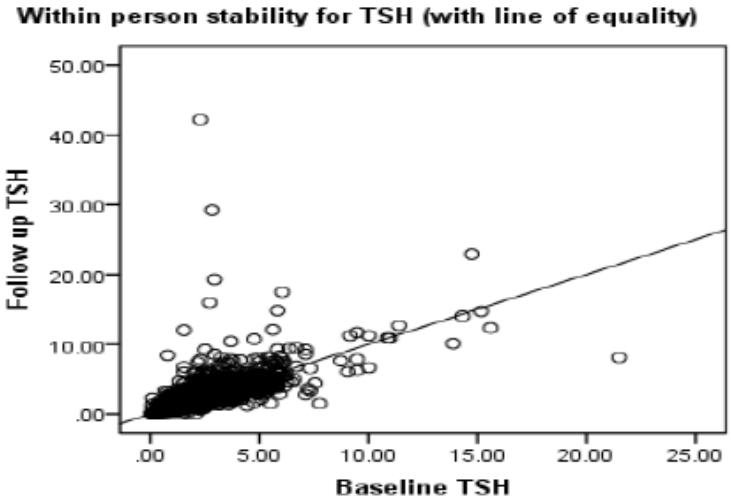


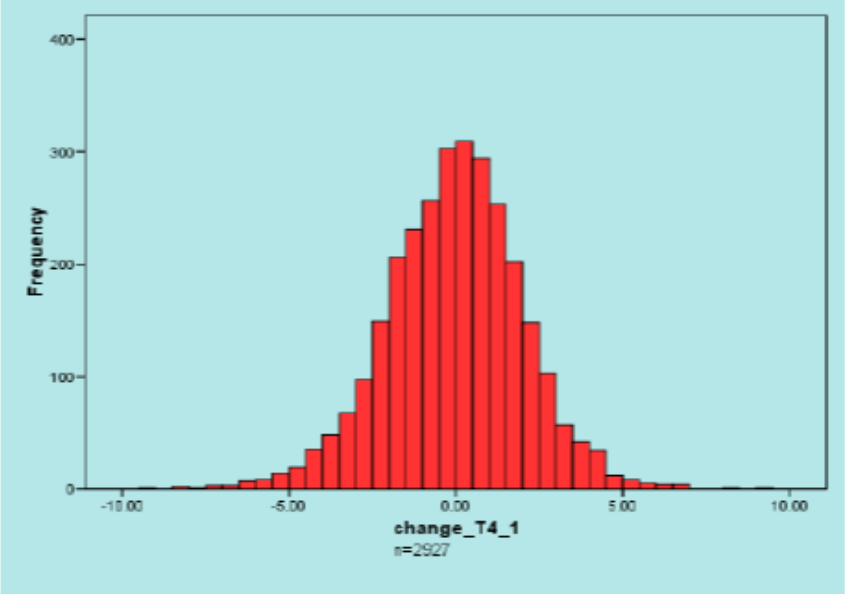
Figure 2: Frequency graphs of change in TSH during the period of follow-up (i.e. approximately 5 years)

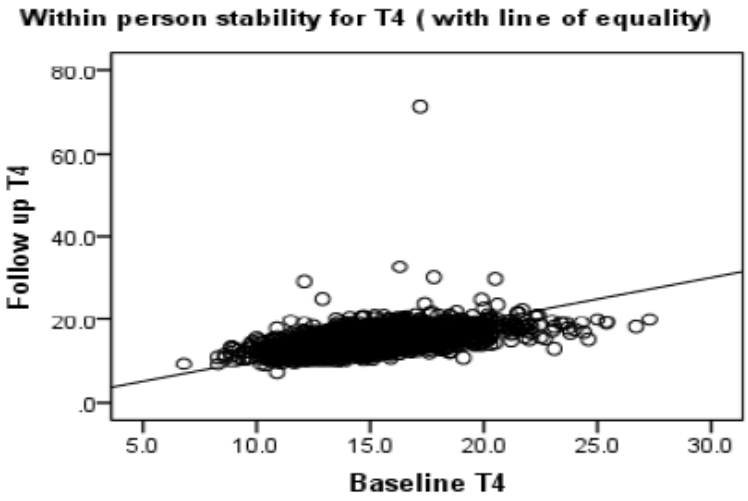




The diagonal line represents identical results at baseline and follow-up testing. Individual participant changes are plotted against this demonstrating clustering around the line of no-change

Figure 3: Frequency graphs of change in FT4 during the period of follow-up (i.e. approximately 5 years)





The diagonal line represents identical results at baseline and follow-up testing. Individual participant changes are plotted against this demonstrating clustering around the line of no-change

Funding source and role of the funder

The study was funded by the National Institute for Health Research (NIHR) School of Primary Care Research. The funders of the study had no role in the study design, data collection, data analyses, data interpretation or writing of the manuscript.

Ethical approval

Ethical approval for the study was obtained from the North Staffordshire Local Research Ethics Committee; reference number: 07/H1204/136, approval date; 19/12/2007 prior to commencement of the research. R&D approval was obtained from the Birmingham and Solihull PCT Consortium (153/1108/P) and Coventry Teaching PCT (COV100907).

Competing interests

None declared

Contributions

The study was designed, and funding were secured by FDRH, LR, JP and DM. LR provided overall supervision of the research. DM undertook day to day management of the study and data collection and LR and DM were responsible for data management and quality assurance. SH provided senior quantitative methodological support for the design of the statistical analysis and undertook all statistical analyses. DM was responsible for descriptive analysis. All authors contributed to data interpretation. LR and OJ wrote the first draft of this paper and all authors were responsible for subsequent critical revision of the manuscript. FDRH is the guarantor and corresponding author for this paper. All authors read and approved the final manuscript. All authors affirm that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Acknowledgements

The authors would like to thank the patients and GPs that participated, the general practice staff that facilitated collection of data and Mr Peter Lewis, Dr Penelope Clark and the staff at the Regional Endocrine Laboratory who conducted the thyroid function testing. The co-investigators for BETS I and/or BETS 2 are Prof Sue Wilson, Prof Jayne Franklyn, Prof Michael Gammage, Mr Roger Holder, Mrs Andrea Roalfe and Dr Helen Pattison. FDRH acknowledges part support from the NIHR School for Primary Care Research (SPCR), the NIHR Collaboration for Leadership in Applied Research in Health and Care (CLARHC) Oxford, and the NIHR Oxford Biomedical Research Centre (BRC).

References

1 Parle JV, Franklyn JA, Cross KW, et al. Prevalence and follow-up of abnormal thyrotrophin (TSH) concentrations in the elderly in the United Kingdom. *Clin Endocrinol.* 1991;**34**(1):77–83.

2 Koutras DA. Subclinical hyperthyroidism. *Thyroid.* 1999;**9**(3):311–315

3 Eggertsen R, Petersen K, Lundberg PA, et al. Screening for thyroid disease in a primary care unit with a thyroid stimulating hormone assay with a low detection limit. *BMJ.* 1988; **297**(6663):1586–1592.

4 UK National Screening Committee. Screening for thyroid disease. https://legacyscreening.phe.org.uk/thyroid (accessed 22/01/18)

5 Surks MI, Ortiz E, Daniels GH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. *JAMA*. 2004; **291**(2):228–238.

6 Vaidya B, Ukoumunne OC, Shuttleworth J, et al. Variability in thyroid function test requests across general practices in south-west England. *Qual Prim Care*. 2013;**21**(3):143–8.

7 The NHS Atlas of Variation in Diagnostic Services. Reducing unwarranted variation to increase value and improve quality, Nov 2013 file:///O:/25.%20Papers/BETS/Right_Care_Diagnostics_Atlas_2013.pdf (Accessed 15/01/18)

8 Allport J, McCahon D, FDR Hobbs, Roberts LM. Why are GPs treating subclinical hypothyroidism? Case note review and GP survey. *Primary Health Care Res Dev,* 2013;**14**(2):175–84

9 Beckett G, Toft AD. First line thyroid function tests - TSH alone is not enough. *Clin Endocrinol* 2003;**58**:20–21

10 Vanderpump MP, Tunbridge WM, French JM, et al. The incidence of thyroid disorders in the community: a twenty-year follow-up of the Whickham Survey. *Clin Endocrinol.* 1995;**43**(1):55–68.

11 Diez JJ, Iglesias P, Burman KD. Spontaneous normalization of thyrotropin concentrations in patients with subclinical hypothyroidism. *J Clin Endocrinol Metab.* 2005;**90**(7):4124–4127.

12 Huber G, Staub JJ, Meier C, et al. Prospective study of the spontaneous course of subclinical hypothyroidism: prognostic value of thyrotropin, thyroid reserve, and thyroid antibodies. *J Clin Endocrinol Metab.* 2002; **87**(7):3221–6.

13 Rosario PW, Carvalho M, Calsolari MR. Natural history of subclinical hypothyroidism with TSH ≤ 10 mIU/l: a prospective study. *Clin Endocrinol* 2016; **84**: 878-881

14 Sawin CT, Geller A, Kaplan MM, Bacharach P, Wilson PW, Hershman JM. Low serum thyrotropin (thyroid-stimulating hormone) in older persons without hyperthyroidism. *Arch Intern Med.* 1991;**151**(1):165–168.

15 Woeber KA. Observations concerning the natural history of subclinical hyperthyroidism. *Thyroid.* 2005;**15**(7):687–691.

16 Wilson S, Parle JV, Roberts LM, et al. Prevalence of subclinical thyroid dysfunction and its relation to socioeconomic deprivation in the elderly: a community-based cross-sectional survey. *J Clin Endocrinol Metab.* 2006;**91**(12):4809–4816.

17 Parle J, Roberts L, Wilson S, et al. A randomized controlled trial of the effect of thyroxine replacement on cognitive function in community-living elderly subjects with subclinical hypothyroidism: the Birmingham Elderly Thyroid study. J Clin Endocrinol Metab. 2010;**95**(8):3623–32.

18 Gammage M D, Parle J V, Holder R L, et al. Association between serum free thyroxine concentration and atrial fibrillation. *Arch Intern Med.* 2007;**167**(9):928–934

19 Roberts LM, Pattison H, Roalfe A, et al. Is subclinical thyroid dysfunction in the elderly associated with depression or cognitive dysfunction? *Ann Intern Med* 2006;**145**(8):573–581.

20 Noble M, Wright G, Dibben C, et al. The English Indices of Deprivation 2004 (revised). http://www.simonpoulter.co.uk/iod/iodpdf/odpm_urbpol_029534.pdf (Accessed 22/01/18)

21 Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. Endocr Rev. 2008;**29**(1):76–131.

22 Kim TH, Kim KW, Ahn HY. Effect of seasonal changes on the transition between subclinical hypothyroid and euthyroid status. J Clin Endocrinol Metab, 2013;**98**(8):3420–9
23 Wong ET, Bradley SG, Schultz AL. Elevations of thyroid-stimulating hormone during acute nonthyroidal illness. Arch Intern Med. 1981;**141**(7):873–5.
24 Diez JJ, Iglesias P. Spontaneous subclinical hypothyroidism in patients older than 55 years: an analysis of natural course and risk factors for the development of overt thyroid failure. J Clin Endocrinol Metab.2004;**89**(10):4890–7.
25 Somwaru LL, Rariy CM, Arnold AM, Cappola AR. The natural history of subclinical hypothyroidism in the elderly: the cardiovascular health study. *J Clin Endocrinol Metab*. 2012;**97**(6):1962–9
26 Surks MI, Boucai L. Age- and race-based serum thyrotropin reference limits. J Clin Endocrinol Metab. 2010;**95**(2):496–502.
27 Chonchol M, Lippi G, Salvagno G. Prevalence of subclinical hypothyroidism in patients with chronic kidney disease. Clin J Am Soc Nephrol, 2008;**3**(5):1296–300